3

Before the
Copyright Royalty Board
Washington, D.C. 20540

In re	)	
	)	
NOTICE AND RECORDKEEPING FOR USE OF SOUND RECORDINGS UNDER STATUTORY LICENSE	)	Docket No. RM 2005-2

SUPPLEMENTAL COMMENTS OF HARVARD RADIO BROADCASTING COMPANY

Harvard Radio Broadcasting Co., Inc. ("WHRB"), respectfully offers these comments in response to the Copyright Royalty Board's supplemental request for written comments on proposed regulations for the delivery and format of records of use of sound recordings under statutory license, 37 C.F.R. Part 270. 70 Fed. Reg. 43364 (July 27, 2005).

I. Introductory Statement

Whatever the merits of these proposals may be as applied to the larger, for-profit webcasters, the proposals for census-type recordkeeping and reporting are unnecessarily burdensome and impractical to be applied to the non-commercial webcasters, 17 U.S.C. § 114(f)(5)(E), whose particular circumstances Congress recognized in the Small Webcaster Settlement Act of 2002, 116 Stat. 2781. The Board would err in attempting to adopt a "one-size-fits-all" approach.

WHRB agrees with the Board that conducting this technical standards-setting procedure through the Library of Congress "does not draw on a reservoir of traditional agency expertise" and is "an undesirable substitute for industry agreement."[1] Furthermore, while the current regulations pertain only to the narrow (and commercially marginal) segment of the music industry known as "webcasting," it is our belief that setting standards for the delivery of electronic royalty-tracking data may play a central role in future rulemaking procedures—whether for music or other media of copyrighted content. In addition, technology moves at a fast pace, and it is quite possible that standards adopted in the current procedure will become outdated in several years, requiring the parties to re-litigate this matter in front of the Board. Therefore, WHRB urges the Board to form a standing committee with expertise in the technical aspects of digital media, including data standards, formats and delivery. Attempts to achieve this result under non-governmental auspices have failed to produce tangible results. See SoundExchange's notification concerning status of settlement discussions filed in Dkt. No. 2005-1 DTRA on July 5, 2005. This committee, to be organized under the principles adapted[2] from the Federal Advisory Committee Act of 1972, as amended, 5 U.S.C. App. 2, could prove a valuable advisor to the Board on rulemaking issues that require technical expertise beyond the Board's "reservoir of traditional ... expertise."

Akin to the music industry, the United States healthcare industry has faced a transition to digital recordkeeping over the past decade. Where the DMCA has required the Library of Congress to promulgate technical standards, the Health Insurance

[1] 70 Fed. Reg. at 43365, -68.

[2] The Act is mandatory only as to agencies as defined in Section 2(a) of the Administrative Procedure Act, as codified to 5 U.S.C. § 551(1).

Portability and Accountability Act of 1996 ("HIPAA") required the Secretary of Health and Human Services to set exacting data and security standards for digital medical health records. The National Committee on Vital and Health Statistics[3]—a public advisory body to the Secretary of Health and Human Services—created a subcommittee on Standards and Security to help the Secretary promulgate these technical standards. By bringing together interested parties along with technical experts, the subcommittee is able to successfully set data standards under the auspices of a governmental agency. WHRB urges the Board to consider this model as an alternative to pushing through technical standards in the absence of industry consensus or cooperation.

WHRB's comments consist of four parts: (1) observations on the burdens associated with statutory licenses and on whom they should be placed with respect to 17 U.S.C. §§ 114(f)(4)(A) and 112(e)(4); (2) specific answers to the factual questions posed by the Board in Section III of the supplemental request; (3) a reiteration that sample-reporting for small, non-commercial webcasters does in fact meet the record-of-use requirements of 17 U.S.C. §§ 114(f)(4)(A) and 112(e)(4); and (4) a short survey of current research on creating efficacious Standard Setting Organizations ("SSO") and why this research concludes that the Copyright Royalty Judges should not be prescribing technical formatting and delivery requirements absent the assistance of a committee comprised of interested parties to the proceedings and technical experts.

II. On Assigning Burdens

[3] See http://www.ncvhs.hhs.gov/ for information on the National Committee on Vital and Health Statistics and http://www.ncvhs.hhs.gov/stdschrg.htm for the mission statement of the NCVHS Subcommittee on Standards and Security.

Assigning burdens between copyright owners and licensees of a statutory license is like assigning blame in a failed marriage: you will generate a lot of arguing without getting any closer to an amicable resolution. From a practical, legal and policy perspective, WHRB believes the correct response is to require fact-based compromise. Both copyright owners (through their agents SoundExchange, Royalty Logic, etc.) and some licensees will be required to make changes in their existing recordkeeping and data processing systems. The difficulty of a compromise is striking the correct balance—as demonstrated by the protracted nature of this rulemaking procedure.

In this section, WHRB will rebut Comments filed May 27, 2005, by SoundExchange concerning the burdens of distributing royalties. WHRB will then briefly outline the various burdens that would be imposed on small, non-commercial webcasters by the April 27, 2005, notice of proposed rulemaking.

In its comments, SoundExchange claims:

> Statutory licensees could be required to provide each copyright owner whose works are transmitted by the service with the direct notice of use under a plain reading of the statutory reporting requirement. As an accommodation to statutory licensees, however, copyright owners and performers—at their own considerable expense—created SoundExchange to handle the collection and distribution of statutory royalties.[4]

WHRB submits that it never was the intent or language of the statutory license provisions to require licensees to directly contact individual copyright owners. According to Congress, the intent of statutory webcasting license was to "create fair and *efficient* licensing mechanisms that address the complex issues facing copyright owners

[4] SoundExchange comments at 3 (May 27, 2005).

and copyright users as a result of the rapid growth of digital audio services."[5] An efficient license is certainly not one in which the licensee is required to track down and contact directly the copyright owner for every sound recording broadcast. Such is not the present practice, and there is no showing it should be for the future. In fact for the small webcasters it is not at all practical. For example, WHRB estimates that it might transmit recordings from 30,000 to 40,000 unique performing artists and from 10,000 to 12,000 different recording labels each year.[6] If Congress were read as intending the station to serve notice on 52,000 copyright owners and artists annually, the statutory webcasting license would not serve its purposes in encouraging new ways of distributing music in a digital environment.

Furthermore, according to SoundExchange, a plain reading of the statute implies that services be required to deliver notices of use directly to each copyright owner under the "statutory reporting requirement," but makes no mention of whether under this "plain reading" licensees would also be required to delivery royalty payments directly to each copyright owner. SoundExchange remains silent on royalty distribution because the statute is quite clear that an agent will be designated to "distribute receipts from the licensing of transmissions in accordance with subsection (f)."[7] It would seem an odd "plain reading" of the statute that would require that notices of use be delivered directly to copyright owners while royalties (destined for these very same parties) be delivered to

---

[5] Conference Report, Joint Explanatory Statement of the Committee of Conference, H.R. Rep. No. 105-796 at 79-80 (1998) (emphasis supplied).

[6] The above estimate does not include the nonfeatured musicians and vocalists who have performed on the sound recordings. Presumably--under SoundExchange's reading of the statute--without a designated agent, WHRB would be tasked with identifying and contacting these parties as well.

[7] 17 U.S.C. § 114(g)(2).

a third-party receiving agent. If SoundExchange finds it troublesome to collate two log files from a single webcaster, how might they collate a check for royalties received directly from a webcaster and 52,000 notices of use delivered to 52,000 independent copyright owners and performing artists?

Since the intent and language of the statute imply that a centralized receiving agent will serve as a clearinghouse for both royalties and notices of use, SoundExchange should not insist that merely creating the SoundExchange organization to "handle the collection and distribution of statutory royalties"[8] counts as a burden or an example of shifting burdens from licensees to copyright owners and performers. .The creation of SoundExchange may be a costly operation, but that is why 17 U.S.C. § 114(g)(3) & (4) explicitly allows for the remuneration of these expenses.

WHRB believes that in the process of setting standards for the formatting and delivery of reports of use, necessary -- but only necessary -- burdens should be shared between copyright owners and licensees. However, when weighing such burdens we do not believe SoundExchange should be credited with relieving licensees of major -- much less unnecessary -- burdens simply through its own existence. The ultimate benefit of SoundExchange existence flows to the licensors for whom SoundExchange is agent in terms of facilitating use of and payment for the digitally recorded music.

In the context of streaming by small webcasters, census-type recordkeeping and reporting are demonstrably neither necessary nor cost-effective. Having operated under statutory licenses for over-the-air broadcasts of musical compositions, WHRB

[8] SoundExchange comments at 3 (May 27, 2005).

understands the trade-offs of necessary burdens which exist between Collective Rights Organizations ("CROs") and licensees. For over thirty years, WHRB has worked with ASCAP and BMI to distribute royalties for the music compositions broadcast by our radio station. The history of our working relationship with these CROs has shown that loose, sample-based reporting—even of a handwritten nature—is sufficient for royalty distribution. Since the royalty rates and numbers generally are quite comparable between the over-the-air license for musical compositions and the statutory license for the transmission of digital sound recordings, WHRB believes the Board must take into account any deviation from our current operating procedures as an unnecessary burden imposed by the new recordkeeping regulations. The following is a quick reminder of the burdens which would be imposed on WHRB should the Board adopt the regulations as described in the April 27, 2005, notice of proposed rulemaking:

1. Keeping logs of all sound recordings transmitted by the radio station. As mentioned in previous comments, WHRB does not keep records of the content of all programs it broadcasts. To do so would require retraining a staff of approximately eighty on-air personnel along with the time required to complete logging while each DJ is broadcasting. For a station staffed by volunteers it would require an unnecessary diversion of effort to supervision that might more constructively be applied to programming.
2. Building and maintaining a system to handle the inputting of logs.
3. Cataloging WHRB's library of 750,000 unique sound recordings. Without a master, digital catalog, it would be impossible for DJs to input logging information while carrying out regular broadcast functions. Cataloging this

number of recordings is a monstrous effort, beyond the capability of a volunteer staff; failure to catalog them would constrict the variety and depth of music programming offered to students in music courses and to the public.

4. Converting our digital logs into the detailed format required by the proposed regulations.
5. Delivering the digital logs to SoundExchange via the complicated processes proposed in the regulations.

WHRB urges the Board to take into account such burdens when adopting final regulations that apportion on a class-specific basis the burdens of only necessary recordkeeping between copyright owners and licensees.

III. Answers to Factual Questions

The following section presents answers to specific factual questions posed by the Board in Section III of 70 Fed. Reg. 43364-43368 (July 27, 2005). These answers are not intended to encompass the issues of whether applying such requirements to student webcasting operations would be necessary or appropriate.

A. Spreadsheets

1. How expensive and time-consuming would it be for a typical noncommercial webcaster on the Internet to compile spreadsheets using Microsoft Excel? Using Corel Quattro Pro?

Since the Board has a second question pertaining to the process of converting a spreadsheet to ASCII format, we will interpret the term

"compile"[9] to mean the process of preparing a spreadsheet program for receiving data and the process of entering data into the spreadsheet.

The first step in using a spreadsheet for logging is to setup the appropriate hardware to run the program. In a real-time, broadcast operation it is advisable to dedicate a single computer to this task instead of sharing a machine which is occupied with other tasks such as streaming music over the internet or communicating via email or instant messenger with listeners. WHRB estimates that it costs $ 600 to purchase a computer and monitor capable of running a spreadsheet program like Excel or Quattro Pro. Best audio-production practices would insist that the physical CPU box not sit in the same room as the broadcast studio. Depending on a station's specific setup and the distance between the CPU box and keyboard/video/mouse ("KVM"), a KVM extender will cost from $ 100 to $ 500. In the case of WHRB, it costs approximately $ 500 in hardware for each remote KVM setup in our main broadcasting studio due to the reinforced concrete surrounding the studio and approximately 100 feet of wires required to reach the electrical closet where the CPU boxes are located. WHRB estimates that it requires five hours of technical labor to setup and install the above hardware system in a broadcast facility. At a below-market-rate of $ 40/hour for IT labor, equipment setup would cost $ 200. The spreadsheet programs—if the station

[9] In computer science, the "compile" term is often reserved for the process of converting a piece of software code from human-readable format into machine-readable format to be interpreted by the operating system. We use the term as defined above, not in its technical form.

does not already own a copy—could be bought with an academic discount for approximately $ 150 each. Therefore, hardware, hardware setup and software for a typical noncommercial webcaster to compile spreadsheets using Microsoft Excel of Quattro Pro would be $ 1050 to $ 1450. Since about half of the campus broadcasters have annual budgets less than $ 9,000, and some as little as $ 500,[10] that is a disproportionate diversion of operating funds.

The major difficulty in compiling spreadsheets—or any census-style logging reports—is data entry. WHRB has conducted in-studio tests with its DJs and determined that entering the fields of data required by the proposed rules into a spreadsheet-like program takes eighty seconds for each sound recording. Assuming that WHRB transmits 55,000[11] sound recordings annually, we estimate that census-style logging will require 1,222 hours of human labor annually. While WHRB is run entirely by student volunteers, if we paid each volunteer the Massachusetts minimum hourly wage of $ 6.75, data entry would cost the station $ 8,250 annually. It is important to note that any census-style reporting scheme—whether by spreadsheet or other computer program—will expend approximately 1,200 hours of data entry by the WHRB staff on an annual basis, and it seems unlikely that the number of annual volunteer-staff hours could be expanded by that amount in practice. Stations with smaller student staffs, would have even less ability to do so.

---

[10] Comments of Inercollegiate Broadcasting System, filed May 2, 2005, in Docket No. RM 2002-1H at 2.

[11] Note this is the number of total—not unique—sound recordings estimated to be transmitted by WHRB annually. Since the logging process will occur in real-time, each individual piece must be entered into the spreadsheet log.

2. *What are the practical difficulties in converting a Microsoft Excel or Corel Quattro Pro spreadsheet into ASCII? How costly is it?*

Between the last round of comments and today's filing, WHRB has been able to test the Microsoft Excel template posted by SoundExchange on its website. WHRB is very impressed with the spreadsheet and its ability to convert a native .xls file to the ASCII variant. While good data practices would suggest that file conversions occur in batch (i.e. SoundExchange converts all native file spreadsheets to ASCII on their end), WHRB believes the proposed template is sufficient to allow the more sophisticated small webcasters to perform the conversions. We estimate that converting each spreadsheet would take one hour of technical labor at a below-market-cost of $ 40. If reporting were done on a quarterly basis, conversion would take four hours annually with a cost of labor of $ 160.

We have one caveat with respect to the template provided by SoundExchange. The Excel template performs the conversion using a macro. Due to the prevalence of computer viruses carried by macros, the default setting on many computers it to prevent their execution. Since WHRB administers our own computers and computer network, it is not difficult for one of our computer technicians to enable macros on an as-needed basis. However, it is possible that in other educational environments where computers are centrally managed, enabling macros could prove troublesome.

We expect that comments from other parties in this proceeding will shed further light on the severity of the macro issue.

3. *What are the kinds of technical support that are typically needed in preparing Microsoft Excel and Corel Quattro Pro spreadsheets and converting them to ASCII? How would that technical support be available to a webcaster and what costs would be involved?*

SoundExchange does a good job at providing instructions with their Excel template to help webcasters properly fill out the spreadsheet and convert the files to ASCII format. The technical support required with the given template would most likely be focused on enabling support for macros as referenced above in § 3(b)(2). Without access to the Quattro Pro spreadsheet, WHRB is unable to determine what types of technical support might be needed to perform successful conversions.

In a typical educational setting, technical support is provided by a centralized University Information Technology group. Depending on a radio station's relationship to the University, the station might be required to pay for technical support on an hourly basis, which might be financially impracticable. In the case of WHRB, technical support is provided by our own internal technology team. While the team operates on a volunteer basis, we estimate its cost to be $ 40 per hour of technical labor.

B. Commercially Available Software

1. *What, if any, commercially available software is available that could be used to compile records of use? Would such software produce records of use that*

are format compatible with SoundExchange's data processing system? What are the costs associated with such software?

Due to the minimal resources available to educationally-affiliated, non-commercial webcasters and their unique operating procedures (i.e. extremely diverse playlists; human DJs who broadcast in real-time; reliance on physical—not digital—source material; etc.), very few pieces of commercial software exist to compile adequate records of use. However, there are several small companies who are trying to fill this niche. The most promising is Spinitron (http://spinitron.com/about/), a two-person outfit which has a beta-product that allows stations to create play logs through the Spinitron website. While Spinitron does not reduce the data-entry task associated with census-based reporting, they do provide a nice interface and centralized database for storage of playlists. They do not currently support the proposed format for reports of use, but the company believes they would be able to modify their system to generate the proper output. Spinitron does not publicly list the prices for their software and services.

C. Report Delivery

1. *What are the average estimated costs of creating and maintaining a Web site for receipt of records of use? What are the security concerns and how may they be addressed? Is there a commercially available Web site software that could perform this task? Is Web site software available that could be adopted from other SoundExchange uses?*

The costs involved for creating and maintained a website for receipt of records of use include two parts: back-end technical infrastructure and front-

end user interface design. Most commercial servers such as Microsoft Windows Server or a Linux Server have built-in support for both an FTP server and web server. The technical back-ends required to receive and store reports of use are virtually identical whether the reports are received via FTP or a website. The major cost difference is on the front-end, user interface. FTP sites require no user interface. Once a user account is established, files are transferred using simple commands or a graphical drag-and-drop file hierarchy tree. In the web environment, a user interface including an account login screen, profile overview and uploading interface are required. WHRB defers to SoundExchange for a cost estimate to build these items. However, WHRB notes that SoundExchange says it is already planning on establishing user accounts for copyright owners on its website, as evidenced by the "Under Construction" page found at the link for "Member Login" on its website.[12] Presumably, many interface items found in the Members section could be adopted for use in creating web accounts for licensees, as well.

WHRB does not believe security concerns underline SoundExchange's refusal to implement a web delivery option. SoundExchange itself supports the use of the FTP protocol to deliver reports of use. Computer security experts widely acknowledge that FTP is an extremely insecure method of data delivery:

> Anybody on the network between the user and the server, with sufficient access to put a network card in promiscuous mode, can easily get your password. As a results, insecure FTP has been

[12] This link can be found at http://www.soundexchange.com/members/login.html

> recognized as one of the largest remaining security holes in many server systems. ... Additionally, unprotected FTP is particularly vulnerable to unintended corruption of data, because it signals end-of-file by closing the TCP connection. Such a closure need not be the result of a hack; any network outage that causes connection closure (e.g. a modem that hangs up) in the middle of a download will have the same effect. Finally, after the user issues a PORT command, anyone can potentially connect to the port and send data, which the FTP client will happily write to disk. This could be used to, for example, substitute a Trojan horse for a downloaded program.[13]

There are methods used to secure FTP, but the proposed regulations do not require these security precautions nor has SoundExchange adopted them in their current implementation. Therefore, WHRB believes a web delivery model would actually be more secure than the current FTP option that SoundExchange finds acceptable.

2. *To what extent can a SoundExchange-hosted Web site reduce costs associated with records of use? Can it assist in organizing and cataloging delivered data and, if so, in what fashion and to what extent?*

WHRB advocates the adoption of a web interface as the preferred method of delivering reports of use because we believe it would aid licensees in properly delivering such reports as are appropriate while helping SoundExchange to automatically tag and organize incoming data.

As an example, BMI utilizes a website interface for collecting Electronic Music Reports. Their system can be viewed at http://emr.bmi.com. It is

[13] Bonachea, Dan and McPeak, Scott, "SafeTP: Transparently Securing FTP Network Services," Report No. UCB/CSD-01-1152, CS Division, EECS Department University of California, Berkeley, February 2001. Report can be found online at: http://www.cs.berkeley.edu/~bonachea/safetp/CSD-01-1152.pdf

interesting that BMI finds the website option as the preferred method of electronic report delivery. The website option creates a useful location for centralizing information about recordkeeping regulations and leads the service through the process of submitting a report. By creating a profile on the site, services would avoid the need to re-enter basic information such as service name, service class, address, etc. each time reports are submitted. WHRB believes the easy-to-use nature of a website will greatly facilitate the ability of webcasters to deliver digital reports of use.

By accepting data through a web-based interface, SoundExchange could better control the quality of reports which enter its data processing system. One of the more difficult parts of the proposed regulations is the proper use of headers containing information about the report of use. The current reporting mechanisms (i.e. FTP, email, CDROM, diskette) only allow one-way interaction—the service must provide information without any feedback or prompting from SoundExchange. On the website, SoundExchange could collect information which is currently required to appear in the header of a report file. Therefore, instead of relying on the service to properly notate the file, SoundExchange would insure proper tagging of all files entering its system.

3. *Could a SoundExchange-hosted Web site be required to provide services with access to prior submitted records of use? For how long?*

Yes, a SoundExchange-hosted website could be required to provide services with access to prior submitted reports. For that matter, a SoundExchange-hosted FTP site, could also be required to provide services with access to prior submitted reports. Since SoundExchange is unlikely to destroy these reports upon receipt—"because a report of use may not be analyzed (through extraction, transformation and loading) by SoundExchange for weeks or months following receipt"[14]—it does not seem difficult for the files to be left accessible in a service's web or FTP directory. Due to issues of staff turnover at educationally-affiliated, non-commercial webcasters, we believe keeping access to twelve months of reports would greatly aid services in preparing future reports of use while enabling services an easier method to make sure current reports are up-to-date.

D. Report Delivery

1. What standing does RLI have to request copies of the reports of use?

The number of sites is largely immaterial, so long as the reporting entities are not required to vary the submissions. Inconsistent formats and requirements would inevitably lead to confusion, delay, and error in reporting, particularly given the relatively short life of a high school or collage generation, i.e., high turnover among volunteers.

2. How expensive and burdensome would it be, on average, for services to provide RLI with records of use in addition to SoundExchange?

[14] SoundExchange comments at 21 (May 27, 2005).

Assuming that RLI accepts reports of use with the common format and delivery specifications as SoundExchange, the major burden for services to deliver duplicate reports would be the time required to deliver the report. We estimate this to be an extra hour of technical labor at a cost of $ 40 per report. This contradicts comments of RLI which indicate that a requirement to deliver duplicate reports "would present no additional burden to the transmission services."

However, in RLI's comments of May 26, 2005, at 2, RLI only includes a physical address and email address for addition to the proposed regulations -- no details on how to access a RLI-hosted FTP site are included. If RLI wishes to have equal rights to reports of use, they must agree to implement all format and delivery regulations. Otherwise, it will be impractical for small webcasters to submit additional reports.

While delivering a single, additional report is not terribly burdensome, WHRB is worried that the number of designated agents might increase over time. Delivering two duplicate reports is doable, but delivering five or ten is not efficient.

3. Must all the format requirements be the same?

Yes, format and delivery requirements must be identical for all reports of use delivered to any designated agent. It is extremely burdensome in human resources to prepare reports with identical data but different formatting specifics and would

inevitably lead to confusion, delay, and errors for the reasons described above. Additionally, delivery requirements should be identical for SoundExchange and RLI if the Board determines that RLI has standing to receive reports of use directly from services.

E. Field Delimiters and Text Indicators

1. *What are the industry standards for use of field delimiters and text delimiters? Should particular ones be specified in the regulations? To what extent is flexibility acceptable in their selection?*

The computer industry has multiple standards for data formatting and data exchange. The digital music industry uses multiple data standards and it is not possible to specify a single "industry standard."

Perhaps the most popular data formatting standard is Extensible Markup Language (XML).[15] The beauty of XML is that it includes a standard formatting language for talking about formatting standards. Therefore, instead of settling on a single data standard, XML allows a file author to specify the format of her data file in a way other computer programs can automatically understand. Applying this concept to the task at hand, XML could obviate the need for the parties to argue over the specific contents of headers, field delimiters, uppercase text, etc. Furthermore, the mere existence of XML demonstrates that much of the current proceeding is merely reinventing already decided concepts in best-practices data

[15] For an excellent overview of XML along with references see: http://en.wikipedia.org/wiki/XML

formatting and delivery.[16] While WHRB would advocate that the Board accept XML data feeds for records of use, we believe that the current proceedings are not adequate for promulgating complex, technical standards such as XML. We reiterate that an advisory board comprised of technical experts and appointees from interested parties working in a collaborative environment would be better suited to adopting these more advanced data formats.

A second major problem in picking a data standard is the presence of a large number of legacy computer systems. It is apparent from the filings of NRBMLC/Salem (May 27, 2005), that ensuring their legacy radio automation and playlisting software interoperates with SoundExchange's system is a major concern. Again, this problem has been tackled by other CROs, such as BMI through their online, Electronic Music Report system.[17]

2. *What problems will be created by allowing the use of commas and quotes as field delimiters and text indicators, respectively? How can such problems, if any, be avoided?*

To clarify our previous comments, WHRB is advocating that SoundExchange accept data in the comma-separated values (CSV) data format. CSV is a widely-supported, ASCII-based format used to exchange data between a large number of disparate computer applications. It is supported directly by

[16] In addition, as XML (and other advanced data formatting schema) become widely adopted, it is likely the parties will need to re-litigate this matter in front of the Board. The difficulty in the current process underscores the need to create a better forum for technical standards-setting under the auspices of the Library of Congress.

[17] Instruction for creating a properly formatted EMR from a variety of radio software programs can be found here: http://emr.bmi.com/CreateEMR.asp?Page=6&From=2

Microsoft Excel. "In legacy systems though (pre-XML), CSV files had indeed become a de facto industry standard."[18]

While the CSV file format is extremely popular, it was never formally documented until April 2005 by Y. Shafranovich from the Network Working Group of the Internet Engineering Task Force (IETF).[19] WHRB has included the full definition for the CSV file format as **Appendix A**. As an overview, the CSV format is characterized by the use of a comma as a field delimiter. Double-quotes can surround fields (but are not required), with their main purpose being to allow commas to appear in textual strings. Please see **Appendix A** for the full specifications.

WHRB is requesting the Board allow properly formatted CSV files as one of the data formats for reports of use to the extent required. As the format is used widely for a variety of data types and is well supported across a range of computer applications, WHRB does not foresee any problems in its circumstances arising from its use.

F. Data Fields

1. *What are the costs/benefits of requiring all data fields to be in upper case characters? Will the SoundExchange data processing system accept lower case characters in a data field and combinations thereof?*

---

[18] Creativyst, Inc., "The Comma Separated Value (CSV) File Format: Create or parse data in this popular pseudo-standard format" 2005. Report accessed online at: http://www.creativyst.com/Doc/Articles/CSV/CSV01.htm

[19] This format specification draft can be found at: http://www.ietf.org/internet-drafts/draft-shafranovich-mime-csv-05.txt

The fact that SoundExchange requests data to be submitted in all upper case characters implies that their matching program does not use case as a determining factor when processing reports. Therefore, it should be extremely simple for SoundExchange to completely ignore the case of submitted data or merely convert the data to upper case characters for processing.

2. What is the industry standard for data fields?

There is no industry standard for representing metadata about music. However, it is widely accepted that data which will either be input or accessed by humans (as opposed to machines) is represented in a mixture of upper and lower case characters. A quick check of any online music service (such as Napster, iTunes, Rhapsody, etc.) will show that the jam band *moe.* is represented as *moe.* and the punk band *M.I.A.* is displayed as *M.I.A.* This implies that metadata for digital music services does indeed maintain the integrity of upper and lower case characters.

To the extent human DJs would be inputting by-hand music metadata, WHRB urges the Board to require SoundExchange to accept data with both upper and lower case characters.

G. Abbreviations

1. What problems, if any, does allowing abbreviations within data fields present to SoundExchange's data processing system? How can these be addressed?

2. *Can a set of rules be developed that permit abbreviations within data fields and, if so, what should these rules be?*

A standard set of abbreviations for music metadata does not exist because artists and performers do not follow a standard set of conventions when naming themselves, their albums or their tracks. For example, the common English abbreviation for the word 'junior' is 'jr.' However, consider the following list of artists or performers which contain the word 'junior' and the preferred spelling of their names:

Preferred Artist Name	**Genre**	**Time Period**
Junior	R&B	70s-90s
Jr	Electronica	90s
jUNIOR	Rock	00s
Dinosaur Jr.	Indie Rock	80s-00s

Given the large number of diverse artist, album and track names, WHRB suggests that the Board adopt flexible regulations which ask services to reports—as best they can—the name of an artist, album or track as it appears on the marketing language which accompanies the sound recording. Services should not be penalized if an attempt is made to accurately capture this data.

3. *What are the burdens and costs associated with the creation and maintenance of a data-base of sound recording titles, album titles, artists' names, etc. by SoundExchange? What should be the functionality of such a database? How could such a database be utilized to reduce the overall costs of reporting records of use?*

The most difficult step in distributing royalties from the statutory webcasting license is correctly matching the sound recording transmitted by a service to an item in SoundExchange's database of known sound recordings and their respective copyright owners. This process would be greatly simplified if the services and SoundExchange were to "speak" to each other in the same language. Instead of squabbling over abbreviations, spellings and capitalization, the services would transmit a unique ID code for each individual sound recording. In the digital music industry, this is known as using a "common metadata ID system."[20]

Obviously, requiring SoundExchange to publicly host (i.e. provide real-time access to) this database is burdensome. However, WHRB envisions an alternate scenario. At regular intervals, SoundExchange will publish a snapshot of its metadata database and ID system. While the database should contain fields such as unique identifier, artist name, album name, track name, and recording label, it should not contain more "proprietary" information such as copyright owner mailing address or bank account information. The public database could be downloaded and used by third-party software developers and music services when constructing their logging systems. For example, Spinitron—the commercial outfit mentioned in §3(B)(1)—could import this database into their playlist server and logging front-end. Therefore, when

[20] Currently, a standard metadata ID system for music does not exist. For example, music services such as Napster, MusicNet, etc. need to manually match their own ID system to systems maintained by external parties in order to communicate effectively about their respective music catalogs.

stations use Spinitron, a large number of their sound recordings would already exist in the database without the need for manual data-entry. In addition, the Spinitron reporting function could be customized to provide reports of use which contain the standardized ID associated with each sound recording. Therefore, the need to match reports of use to SoundExchange's database would be obviated.

Making this database publicly available would not be a major burden for SoundExchange. The database itself is a by-product of SoundExchange's routine operating procedure and will require no extra effort to compile. The resources required to publish the database could easily be provided by an outfit such as freedb.org—an organization which currently mirrors a large music metadata database across multiple servers. Conceivably, SoundExchange could claim that this database constitutes confidential and proprietary information of the company. WHRB thinks this is a fallacious argument. SoundExchange is building this database from the raw data provided by the services. While information such as mailing addresses and bank account information are not public (and might indeed be proprietary property of SoundExchange), the data WHRB is requesting should most certainly belong in the public domain.

Setting up such a system would have two major cost-saving features: (1) SoundExchange would expend exponentially less resources while achieving nearly 100 percent accuracy in matching; and (2) services would be saved the

burden of cataloging by hand their entire music libraries. The caveat remaining as to the smaller webcasters is their ability to manually handle such a complicated system for census-type reporting.

While SoundExchange does not publicly make available its success rate in matching reports of use, industry average for purely lexical matching of disparate music metadata sets is roughly 70 %.[21] This would imply that even with census-style reporting, SoundExchange could be missing up to 30% of each report due to lexical mismatches. By sharing a common ID system with the services, SoundExchange's match rate success could be raised to almost 100 % accuracy.[22]

Improving matching accuracy is extremely important because it ensures more artists and copyright owners receive their royalties. Previously, SoundExchange has rejected the adoption of sample-based reporting—a major request of non-commercial webcasters—on the basis that it "may results in the non-payment of royalties to over thirty percent of the performers entitled to such royalties."[23] Even ignoring the cost savings of implementing a standardized ID system, the increased efficiency for royalty distribution is grounds to adopt WHRB's proposal. Therefore, for the same reasons

---

[21] This figure comes from research conducted by MediaUnbound, Inc. in matching an internal music metadata database to multiple external metadata sources on a purely lexical basis. MediaUnbound has over five years of industry experience conducting this type of work. More information on MediaUnbound can be found at http://www.mediaunbound.com or by contacting the undersigned, who is a principal.

[22] A second method for improving match rate success is to utilize human brute-force labor to manually fix unmatched items. This method is resource intensive and costly.

[23] SoundExchange comments at 9 (May 27, 2005).

SoundExchange advocates that the Board mandate census-style reporting, WHRB advocates that the Board mandate SoundExchange to make publicly available a comprehensive, universal database to identify sound recordings. As SoundExchange pleads, "excluding so many performers from the royalties to which they are entitled seems antithetical to the intent of the statutory license, and SoundExchange respectfully requests that any requirements that result in the exclusion of so many performers be corrected with the adoption of final regulations."[24]

The second major advantage to adopting a standardized ID system is the savings it will provide the services when cataloging their music libraries. As WHRB has previously commented, we have a music library of physical media estimated to contain 750,000 sound recordings. To facilitate real-time logging, it is desirable to have a digital catalog containing information about this music such as artist name, album name, track name, media type, recording label, etc. To enter all of this data by-hand is extremely time-consuming and burdensome. However, by pooling submissions from all of the reporting music services into a single, standard database, webcasters would be able to share the burden of the massive data-entry task. The standard database will improve the accuracy of reports of use and reduce the amount of time it takes for non-commercial webcasters to prepare these reports. Finally, as WHRB has mentioned previously, we believe the creation of a standard metadata

[24] Ibid.

database will have positive and wide-reaching effects on the fledgling digital music industry.

IV. On the need for sample-based reporting

Census-based reporting is the largest, single burden associated with recordkeeping that can be imposed on smaller, non-commercial webcasters. It is not at all clear why SoundExchange or the Board should choose to implement this burdensome new scheme of census-reporting, when a more practical and reasonable scheme has performed satisfactorily for over three decades for BMI and ASCAP. The royalty rates and numbers generally are quite comparable. ASCAP has been able to distribute royalties with music logs—even handwritten—of one week annually from a sampling of small stations. BMI has been able to satisfy its needs with a seventy-two hour music log for each station annually. SoundExchange has failed to demonstrate why anything more would be reasonable.

The statute is vague on the subject of recordkeeping. Sections 114(f)(4)(A) and 112(e)(4) ask only that the Board "establish requirements by which copyright owners may receive reasonable notice of the use of their sound recordings." WHRB believes that sample-based reporting for a subset of webcasters—small, non-commercial webcasters—meets this test for reasonableness on two grounds.

First, copyright owners have historically been able to distribute royalties for the public performance of copyrighted music compositions by using loose, sample-based reporting from smaller broadcast entities. If these copyright owners have been able to

effectively distribute royalties for over 30 years using this scheme, we must conclude the practice constitutes a "reasonable" notice of use. WHRB agrees with SoundExchange that an *ideal* requirement for notice of use would be census-based reporting. However, the statute is clear that we must only meet the test for reasonableness, even if more ideal forms of reporting may exist.

Secondly, a quick economic analysis shows that loose, sample-based reporting is reasonable. Ignoring the cost—in resources and time—for building and maintaining computer systems for generating and delivering electronic reports, WHRB estimates that the average educationally-affiliated, non-commercial webcaster would spend 1,200 hours (or a cost of labor of $ 8,250) annually compiling census-based reports of use, if such additional volunteer hours were available. The bulk of non-commercial webcasters will pay $ 250 per station per year in royalties. It is not reasonable to spend thirty-three times the annual royalty fee on generating reports of use. Certainly, if SoundExchange were to spend thirty-three times the royalties collected on the royalty distribution process, we would deem this unreasonable—and bad business. The same should be true when analyzing the impact on non-commercial webcasters.

WHRB believes loose, sample based reporting meets the test of reasonableness for smaller, non-commercial webcasters. We ask the Board to implement final regulations along the lines of the recordkeeping applied by ASCAP and BMI to this class of stations.

V. <u>On Standards-Setting Organizations</u>

The current proceeding before the Board is comprised of two parts: (1) policy and legal decisions about the reasonableness of certain recordkeeping practices (i.e. sample vs. census; electronic vs. handwritten); and (2) standard setting for a format to exchange data between the parties. While we believe that the Board (and the Copyright Office, previously) have tried very hard to understand the nuances of standard setting, we do not think the current forum is conducive to proper setting of technical standards.

Certainly, the preferred method in setting a standard would be for the parties to engage directly in private negotiation and settlement. However, for a broad consensus, the parties need to engage in parallel negotiations. WHRB believes that the Board should facilitate this type of collaborative-based, standard-setting forum under the auspices of the Library of Congress.

The history of this proceeding demonstrates the value of collaborative discussion about data formats and standards. Most of the written comments in this docket are fairly contentious and show little willingness to compromise by any of the parties. Academic research on standard setting would predict this is the case when a single arbiter (the Board) has unilateral power to adopt a final standard.[25] However, on October 8, 2002, the Office convened the parties at the Library of Congress for a status conference on the proceedings. WHRB believes that this event made the most forward progress towards a data standard. While most parties in attendance were represented by legal counsel, the

[25] For relevant research on the organizational structures of SSOs, see Christopher T. Marsden, "Cyberlaw and International Political Economy: Towards Regulation of the Global Information Society, 201 L. Rev. Mich. St. U. Det. C.L. 355, 358-59 (2001) and Chiao, Benjamin Hak-Fung, Josh Lerner and Jean Tirole, "The Rules of Standard Setting Organization: An Empirical Analysis", Working Paper, December 7, 2004.

services were able to engage SoundExchange in meaningful dialog about the need for certain features in preparing reports of use. For example, WHRB was able to demonstrate how a Microsoft Excel file could be used to facilitate the creation of a report of use while fielding questions from both SoundExchange and representatives from the Copyright Office. While final decision-making power still rested with the Office, this status conference was the closest to a collaborative discussion on data standards the parties have engaged in and brought the most compromise to-date.

When faced with the need to promulgate technical data standards, other United States Government agencies have responded by creating advisory forums comprised of interested parties and technical experts to engage in collaborative discussion on standard setting. As mentioned previously, the Secretary of Health and Human Services utilizes a subcommittee on Standards and Security when setting complicated standards for electronic medical health records. These meetings are open to the public and transcripts are made available via the website of the National Committee on Vital and Health Statistics.[26] WHRB thinks this model should be emulated by the Board in the current proceeding and future rulemaking processes which require technical expertise and standards.

II. Prayer

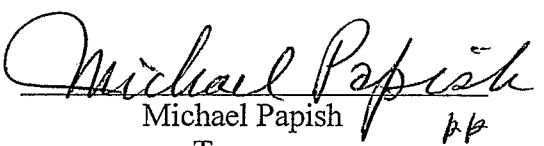
WHRB urges the Board to be mindful of the differences between educational webcasters and large, commercial webcasters when promulgating final regulations

[26] Notes from one such meeting can be found here: http://www.ncvhs.hhs.gov/050113ag.htm

governing recordkeeping for use of digital sound recordings and reporting and to adopt regulations appropriate to small, noncommercial webcasters. As to these stations the Board should reject application of census-type recordkeeping and reporting as unnecessary, impractical, and not cost-effective. In addition, WHRB asks the Board to move the process for setting standards forward by creating a forum in which collaborative discussion of these technical issues can occur.

Respectfully submitted,

HARVARD RADIO BROADCASTING CO., INC.

by 

Michael Papish
Treasurer
Station WHRB (FM)
389 Harvard Street
Cambridge, Massachusetts 02138

Of counsel:[27]

William Malone
Miller and Van Eaton, P.L.L.C.
1155 Connecticut Ave., N.W., # 1000
Washington, D.C. 20036-4320

4122\03\00111505.DOC

[27] Required to be served under P.L. 89-332, 5 U.S.C. § 500(f).

Appendix A

The following is the "Definition of the CSV format" as found in the Internet-Draft by Y. Shafranovich from the Network Working Group of the Internet Engineering Task Force (IETF). The full draft can be found online at: http://www.ietf.org/internet-drafts/draft-shafranovich-mime-csv-05.txt

Definition of the CSV format

While there are various specifications and implementations for the CSV format (for ex. [4], [5], [6] and [7]), no formal specification exists which causes a wide variety of interpretations for CSV files. This section seeks to document the format that seems to be followed by most implementations:

1. Each record is located on a separate line delimited by a line break (CRLF). For example:

   aaa,bbb,ccc CRLF
   zzz,yyy,xxx CRLF

2. The last record in the file may or may not have an ending line break. For example:

   aaa,bbb,ccc CRLF
   zzz,yyy,xxx

3. There maybe an optional header line appearing as the first line of the file with the same format as normal record lines. This header will contain names corresponding to the fields in the file and should contain the same number of fields as the records in the rest of the file (the presence or absence of the header line should be indicated via the optional "header" parameter of this MIME type). For example:

   field_name,field_name,field_name CRLF
   aaa,bbb,ccc CRLF
   zzz,yyy,xxx CRLF

4. Within the header and each record there may be one or more fields, separated by commas. Each line should contain the same number of fields throughout the file. The last field in the record may not be followed by a comma. For example:

   aaa,bbb,ccc

5. Each field may or may not be enclosed in double quotes (however some programs such as Microsoft Excel do not use double quotes at all). If fields are not enclosed with double quotes, then double quotes may not appear inside the fields. For example:

   "aaa","bbb","ccc" CRLF
   zzz,yyy,xxx

6. Field containing line breaks (CRLF), double quotes and commas should be enclosed in double-quotes. For example:

```
"aaa","b CRLF
bb","ccc" CRLF
zzz,yyy,xxx
```

7. If double-quotes are used to enclosed fields, then a double-quote appearing inside a field must be escaped by preceding it with another double quote. For example:

```
"aaa","b""bb","ccc"
```

The ABNF grammar [2] appears as follows:

```
file = [header CRLF] record *(CRLF record) [CRLF]

header = name *(COMMA name)

record = field *(COMMA field)

name = field

field = (escaped / non-escaped)

escaped = DQUOTE *(TEXTDATA / COMMA / CR / LF / 2DQUOTE) DQUOTE

non-escaped = *TEXTDATA

COMMA = %x2C

CR = %x0D ;as per section 6.1 of RFC 2234 [2]

DQUOTE =  %x22 ;as per section 6.1 of RFC 2234 [2]

LF = %x0A ;as per section 6.1 of RFC 2234 [2]

CRLF = CR LF ;as per section 6.1 of RFC 2234 [2]

TEXTDATA =  %x20-21 / %x23-2B / %x2D-7E
```

Certificate of Service

I hereby certify that I have caused to be mailed this day a copy of the foregoing to:

Gary R. Greenstein
General Counsel
SoundExchange, Inc.
1330 Connecticut Avenue, NW
Suite 330
Washington, D.C. 20036

William Malone

Washington, D.C.
August 26, 2005